Casemix Service
HRG4

The Casemix Service

HRG4 Reference Cost Grouper – Guide to File Preparation

Issue Date: 28 April 2008

CONTENTS

DICTIONARY AND GLOSSARY 3
1. INTRODUCTION 4
2. OVERVIEW OF FILE PROCESSING 5
3. ADMITTED PATIENT CARE 8
4. NON-ADMITTED PATIENTS 14
5. EMERGENCY MEDICINE 16
6. ADULT CRITICAL CARE 17

Dictionary and Glossary

The following definitions and abbreviations are provided to assist users to understand some key terms and acronyms. Where definitions are provided they have been stated in a context relating to the use of HRGs (Healthcare Resource Groups).

Term	Expansion/explanation
CART	Classification and Regression Trees.
Casemix classification	Classification of people or treatment episodes into groups, using characteristics associated with the condition, treatment or outcome that can be used to predict need, resource use or outcome.
CC	Complications and Comorbidities.
CDS	Commissioning Data Set.
CV	Coefficient of variation – a measure of the amount of variation within a group of values.
DH	Department of Health.
ERP	Expert Reference Panel.
EWG	Expert Working Group.
HRG	Healthcare Resource Group – groupings of treatment episodes which are similar in resource use and clinical response.
ICD-10	International Classification of Disease and Related Health Problems. Internationally defined classification of disease, managed by the World Health Organisation – 10th Revision.
Iso-resource	Similar in resource use.
LOS	Length Of Stay – Duration of the hospital stay from admission to discharge.
IC	Information Centre for Health and Social Care
OPCS	The system of codes used to record interventions and procedures. This was originally developed by the Office of Population Censuses and Surveys (superseded by the UK Statistics Authority).
PbR	Payment by Results. The financial system providing a transparent, rules-based system for paying trusts where payment is linked to activity and adjusted for casemix.
PCT	Primary Care Trust.
RIV	Reduction in Variance. This is a measure of how much variation is explained by the HRGs. The aim with HRGs is to maximise the RIV.

1. Introduction

This document assists users of the HRG4 Reference Cost Grouper ('the grouper') in preparing data for processing within the application.

Activity data from existing healthcare systems will be prepared and processed as part of the Reference Costs submission to the DH for the year 2007/08. This document identifies the data items required and provides information about the data preparation necessary for successful processing.

This document is not intended as a user guide for the grouper. A help file, which gives full instructions, is incorporated into the grouper software package; this can be downloaded from the Casemix website.

1.1. Why is the Reference Cost Grouper Required?

The Reference Cost Grouper has been developed to:

- Underpin Payment by Results and support the introduction of HRG4
- Prepare data for Reference Costs submission for 2007/08
- Provide HRG4 assignment to activity reflecting changes in:
    a. clinical practice and costs
    b. introduction of HRGs to new clinical areas
    c. introduction of 'setting independence'
- Incorporate the introduction of:
    a. unbundling
    b. improved complication and comorbidity splits
    c. improved identification and classification of procedures using updated OPCS codes
- Enable more accurate analysis of healthcare activity within the service

Further information on Payment by Results can be found at:
http://www.dh.gov.uk/PolicyAndGuidance/OrganisationPolicy/FinanceAndPlanning/NHSFinancialReforms/fs/en

Although this document contains references to adjustments to be made to data for reference costing, a full description of which records to include in, and exclude from, the DH PbR Reference Costs submission is beyond the scope of this document. Further information on Reference Cost data collection can be found at:
http://www.dh.gov.uk/PolicyAndGuidance/OrganisationPolicy/FinanceAndPlanning/NHSReferenceCostsDataCollection/fs/en

2. Overview of File Processing

2.1. Data Input

The HRG4 Reference Cost Grouper allows activity data to be processed into standard groupings of clinically similar treatments which use comparable levels of healthcare resource.

To achieve this it uses the HRG4 definitions for the following activity types:

- Admitted Patient Care
- Non-Admitted Patients
- Emergency Medicine
- Adult Critical Care

Because of differences in the way HRG4 definitions are applied to the above activity types and differences in the data required to correctly assign the HRG, the Reference Cost Grouper only processes a data file for a single activity type at one time.

The activity type must be specified by the user within the grouper prior to the input data file being processed. This is done by choosing the required definition database file from the drop down list on the main screen at start up. See the grouper help for further details.

The software accepts data as a comma-separated text file using the standard ASCII character set.

The data items required for the purpose of Reference Costs submission are largely a sub-set of the data items used for recording Hospital Episode Statistics (HES). These in turn are either derived or taken directly from the CDS data file definitions. This should ensure that the data are readily available from sources that are routinely collected.

The principal data sources from which data for HRG4 groupings are derived are:

- Admitted Patient Care Commissioning data set (CDS)
- Outpatient Attendance CDS
- Accident and Emergency CDS (for Emergency and Urgent Care HRGs)
- Adult Critical Care Minimum Data Set (MDS)

These data are already recorded in many existing Patient Administration Systems (PAS); however, it is likely that data required to support the derivation of HRGs for Chemotherapy and Radiotherapy treatments are not currently being captured in a local PAS. Where these data are only collected in specialist departmental systems, local arrangements will be needed to ensure that the data can be incorporated into admitted care or outpatient datasets for HRG grouping following the guidance in the relevant chapter below.

2.2. Data Preparation

The data items required differ for each activity type and extracts from each proprietary PAS may also differ. It is therefore difficult to fully automate the data input process without some user intervention in the form of data preparation. This is covered in more detail for each activity data type in the chapters below.

In general, the user must provide the Reference Cost Grouper with information regarding which data item in their own input file maps to the required data field for the relevant activity type definition file. The software achieves this by using a Record Definition File (RDF) to interpret the input data file. This cross references each of the required activity data fields to the mapped field position within the user's data file.

2.3. Data Validation

In order to accurately assign an HRG to activity data, the software will ensure that the data are complete, valid and within expected value ranges. The software will apply three stages of validation to the data during a processing run.

The three stages are:

- Field content within record
- Cross validation of episodes within spell
- Grouping Logic (assignment of flag values)

Where the Reference Cost Grouper has not been able to resolve the assignment of an HRG definition to the data contained in an input record, a code of UZ01Z is returned in the output record signifying that the data is invalid for grouping.

If there are errors in the input data these will be reported in the Data Quality Report but processing will not be halted. There can be more than one reason why the assignment was not resolved and so this report may contain several records, all of which need to be reviewed to identify the underlying problems.

In order to relate invalid records to source systems it may be valuable to add a unique key field to each record.

2.4. Approaches and Sites

Within the HRG4 definitions there are instances where the assignment of an HRG to a procedure code is dependent on the presence of other codes, indicating either the approach that was taken or the site of treatment.

Such codes should always be *subsequent* to the procedure to which they refer - this should be taken to mean 'after and adjacent to', unless otherwise specifically stated.

Care should be taken that coding practice reflects this; the software will not apply a single site code to a group of procedures, and neither will it apply an approach code at the end of a record to all of the procedures in that record. Site codes and approach codes are only applied to the procedure that directly precedes them in the input record.

An example is given below:

Episode No	Primary Procedure	Procedure 2	Procedure 3	Procedure 4
1	J104	Z393	J118	
2	J106	J104	Z393	
3	J106	Z393	J104	Z393

In the above record sequence it should be noted that in episode no. 2 the site recorded as Z393 “Portal vein” is taken to refer only to the J104 angioplasty procedure and is not applied to the J106 thrombolysis procedure. If both procedures are performed on the portal vein then the record must be recorded as shown in episode no. 3.

2.5. Data Output

Given a correctly configured RDF and user data file, then the software package can process the data against the chosen HRG4 definitions. The assigned core HRG and any additional information is appended to each record and this is output as a flat file in CSV format with quoted alphanumeric fields. This will allow ease of import into MS Excel, Access or other applications.

The filename and location can be chosen using the Process File dialog box. The output file will have the following features:

- A single output row for each row from the input file.
- Each column of the input file will be output in the same location in the output file.
- The relevant grouping and activity analysis data will be appended to the end of each record.

The data to be appended is specific to each activity type and is covered in the relevant chapters below.

3. Admitted Patient Care

3.1. Overview

Table 1 below shows the data items required by the Reference Cost Grouper.

Table 1.

Field Name	Grouper Description	Format	Validation	Notes / Derivation
PROCODET	Organisation Code	text	Not Null	Local provider code possibly the NACS code (an(5))
PROVSPNO	Hospital Provider Spell Number	text	Not Null	Local Spell Identifier
EPIORDER	Episode Number	number	1 – 87. Distinct within Provider Spell.	File must be ordered by PROCODE, PROVSPNO and EPIORDER
STARTAGE	StartAge	number	0 – 130. Increase between episodes relative to episode duration.	Start Date(Episode) - Person Birth Date (integer whole years)
SEX	Sex	text	0,1,2 or 9. Identical value for all episodes within Provider Spell.	Must be a valid national code
CLASSPAT	Patient Classification	number	Not Null	Must be a valid national code
ADMISORC	Source Of Admission (Hospital Provider Spell)	text	Not Null. Identical value for all episodes within Provider Spell.	Must be a valid national code
ADMIMETH	Admission Method (Hospital Provider Spell)	text	Not Null. Identical value for all episodes within Provider Spell.	Must be a valid national code
DISDEST	Discharge Destination (Hospital Provider Spell)	text	Not Null. Identical value for all episodes within Provider Spell.	Must be a valid national code
DISMETH	Discharge Method (Hospital Provider Spell)	text	Not Null. Identical for all episodes within Provider Spell.	Must be a valid national code
EPIDUR	Episode Duration	number	0-99999	Whole days
MAINSPEF	Main Specialty Code	text	Not Null	Must be a valid national code
TRETSPEF	Treatment Function Code	text	Not Null	Must be a valid national code
NEOCARE	Neonatal Level of Care	text	Not Null	Must be a valid national code (0 if not relevant)
OPER_01	Primary Procedure (OPCS)	text	Null or valid OPCS Code without decimal point.	Dots are not accepted. For example, use F343 rather than F34.3. See note at foot of table.

OPER_02 - 12	Additional procedures	text	As above.	
DIAG_01	Primary Diagnosis (ICD)	text	Not Null. Valid ICD code excluding decimal point character and any dagger/asterisk characters.	Dots are not accepted and any dagger/asterisk codes must be removed. For example, use A170 rather than A17.0† or A17.0D.
DIAG_02 - 14	Additional diagnoses	text	Null or valid ICD-10 Code	See above.
CRITICALCAREDAYS	CRITICALCAREDAYS	Number	Blank or 0-99999. Non-numeric characters treated as zero.	Number of Critical Care days. Blank or zero where the user has already removed Critical Care days from the data or where there are no Critical Care days.

Note: the Record Definition File 'extract' facility can be used to remove decimal points from clinical codes during file processing. See the grouper user manual for details.

3.2. Record Preparation

Organisation Code

This data item can be any code used locally to identify the organisation. We recommend using the three or five character NACS code.

Hospital Provider Spell Number and Episode Number

The Reference Cost Grouper provides data output based on HRGs assigned at both an Episode and Spell level.

It is not necessary to sort the data before processing; the grouper sorts the data automatically during processing.

All records are treated as Finished Consultant Episodes (FCEs). It is the user's responsibility to ensure that the data file contains only those records required for the purpose of Reference Costs submission.

Where the source data does not include data items which are directly equivalent to Hospital Provider Spell Number and Episode Number (or they cannot be extracted from the source systems) then some form of proxy must be included in these fields within the submitted file.

For Episode Number a simple incremental count may be used.

There are several ways to create a proxy to spell number, examples are:

1. Concatenate local patient identifier, Admission Date and Admission Time to create a proxy spell number. **Important:** When concatenating data, do not include any spaces

between data items. Admission Time is suggested in order to identify separate spells starting on the same date.

2. Where unique episode identifiers are assigned then use the first episode identifier within the spell as the spell number.

StartAge

This is the patient's age at the start of the consultant episode. The calculation is:
Start Date (Episode) – patient's date of birth

The age should be an integer value of whole years rounded down. Age is required as a determinant for several HRG assignments the validation procedures will return a UZ01Z code where age is not present or where it is not a value between 0 and 130. The user must therefore extract the correct data from their source data files and calculate the age value for each record of the data input file.

N.B. the validation process will check each FCE record and will not accept the AGE field appearing only within the initial FCE of a Spell.

Sex

Sex can be a value 0, 1, 2, 9 as per ISO 5218. However, where sex is a determinant of the HRG then sex must be a value either 1 or 2 indicating male or female for the correct HRG to be assigned.

Patient Classification

Patient Classification must be a value 1, 2, 3, 4 or 5 as per the national codes. It must be the same value for all episodes within a Hospital Provider Spell.

Source of Admission (Hospital Provider Spell)

This is the national code for the source of admission for the Hospital Provider Spell. The Reference Cost Grouper will check all of the episodes within a spell in order from the earliest FCE forward. It will take the first available value from this field and validate it against the national codes. The software will ignore null values, but users should be aware that a value outside of the national codes will cause a UZ01Z code to be assigned as the core HRG.

Admission Method (Hospital Provider Spell)

This is the national code for the method of admission for the Hospital Provider Spell. The Reference Cost Grouper will check all of the episodes within a spell in order from the earliest FCE forward. It will take the first available value from this field and validate it against the

national codes. The software will ignore null values, but users should be aware that a value outside of the national codes will cause a UZ01Z code to be assigned as the core HRG.

Discharge Destination (Hospital Provider Spell)

For finished spells only, this is the national code for the discharge destination.

The Reference Cost Grouper will obtain the value of this field from the last episode in the spell and validate it against the national codes. A value outside of the national codes will cause a UZ01Z code to be assigned as the core HRG.

Discharge Method (Hospital Provider Spell)

For finished spells only, this is the national code for the discharge method of the patient.

The Reference Cost Grouper will obtain the value of this field from the last episode in the spell and validate it against the national codes. A value outside of the national codes will cause a UZ01Z code to be assigned as the core HRG.

Episode Duration

This is a derived data item which the user must calculate prior to the record being submitted to the Reference Cost Grouper. The basic calculation is:
End Date(Episode) – Start Date(Episode)
To avoid double counting for reference costs submission purposes, both critical care and rehabilitation activity should be subtracted from the Episode and Spell Duration.

To Exclude Adult Critical Care Activity:

For the purpose of Reference Costs, subtract from the Episode Duration the number of days within the FCE that the patient has spent in Adult Critical Care. It is possible for either the Critical Care Start Date or the Critical Care Discharge Date to be different from the Episode Start Date or the Episode End Date so care must be taken in calculating how many days within the FCE have been spent in critical care.

The calculation for the number of critical care days is:
(The earlier of CC Discharge Date or End Date (Episode)) – (the later of CC Start Date or Start Date (Episode))

N.B. Episode Duration should be returned as the number of whole days, rounded down. However, where this would give zero or a negative figure then Episode Duration should be recorded as 1 (one) day.

Automatic Exclusion of Adult Critical Care Days

The grouper application provides a facility to remove critical care days as part of the grouping process.

To use this facility, the input field CRITICALCAREDAYS for each FCE record should be populated with the number of Adult Critical Care days that are to be removed from that FCE.

If Adult Critical Care days have already been removed from the data or there are no Adult Critical Care days to be removed, the input field CRITICALCAREDAYS should either be left blank or populated with zeroes.

Treatment of Rehabilitation and/or Specialist Palliative Care Activity

To avoid double counting activity, episode and spell duration should be adjusted for Rehabilitation and Specialist Palliative Care following grouping. This is because Rehabilitation and Specialist Palliative Care admitted patient activity will generate unbundled per diem HRGs assuming accurate recording of classification codes.
Care should be taken to ensure that the HRGs for the core care do not change in line with the adjusted episode and spell duration.

Main Specialty Code

This is the main specialty in which the consultant is contracted or recognised, it should be a valid code from the Main Specialty Codes table.

Treatment Function Code

This is the treatment function under which the patient is treated. It may be the same as the main specialty code or it may be a valid code from the Treatment Function Codes table.

Neonatal Level of Care

Neonatal Level of Care is used by the HRG4 Reference Cost Grouper to allocate the HRG for Neonatal Critical Care Retrieval (XA06Z).

When a patient is not neonatal, a value of 0 should be supplied, even for adults.

Primary Procedure

The grouper application relies on interventions and procedures being recorded as OPCS codes.

Additional Procedures

Some of the grouper processing uses predefined combinations of codes to simplify the processing logic. These may be combinations of procedure codes referring to approaches used, or the site to which the procedure was applied. These should be recorded as OPCS codes. The grouper application will take all of the procedure codes into account when assigning HRGs.

Primary Diagnosis

This is the clinical classification of the patient's diagnosis; the Reference Cost Grouper only recognises the ICD-10 coding system.

Additional Diagnoses

Where the Reference Cost Grouper cannot assign an HRG based on the procedures within an FCE record, it will attempt to assign an HRG based on diagnosis. In doing so it only takes the primary diagnosis and the secondary diagnosis in the most significant position into account.

However, some HRG assignments based on procedure codes are dependent on the diagnosis recorded and in some instances the position of the diagnosis within the record is irrelevant. The Reference Cost Grouper therefore accepts all of the diagnoses codes recorded in the CDS record.

3.3. File Preparation

The Reference Cost Grouper will process all records in the data input file and treat them as FCE records. It will treat all records with identical values for the Organisation Code and Hospital Provider Spell Number as belonging to the same Spell and will perform spell based processes on the total content.

Non-Admitted Patients

3.4. Overview

Table 2 below shows the data items required by the Reference Cost Grouper for Non-admitted Patients, i.e. outpatient attendances or attendances by patients for nursing care on a ward (ward attenders). Since April 1st 2005 the data source for both of these has been the Outpatient Attendance CDS.

Table 2.

Field Name	Description	Format	Validation	Notes / Derivation
AGE	Age	number	0 - 130	Appointment Date - Person Birth Date (integer whole years)
SEX	Sex	text	0,1,2 or 9	Must be a valid national code
MAINSPEF	Main Specialty Code	text	Not Null	Must be a valid national code
TRETSPEF	Treatment Function Code	text	Not Null	Must be a valid national code
FIRSTATT	First Attendance	text	Not Null	Must be a valid national code
OPER_01	Primary Procedure (OPCS)	text	Null or valid OPCS Code	Dots are not accepted. For example, use F343 rather than F34.3.
OPER_02 - 12	Additional procedures	text	Null or valid OPCS Code	See above.

3.5. Record Preparation

Age

This is the patient's age at the time of the attendance. It is therefore:
Attendance Date – Birth Date

The user must therefore extract the patient's Birth Date from their source data, calculate the age value and place it in each record of the data input file.

The age should be an integer value of whole years (rounded down).

Age is required as a determinant for several HRG assignments. The validation procedures will therefore return a UZ01Z code if age is not present or is not a value between 0 and 130.

Sex

Sex can be a value 0, 1, 2, 9 as per ISO 5218. However, where sex is a determinant of the HRG then sex must be a value either 1 or 2 indicating male or female for the correct HRG to be assigned.

Main Specialty Code

This is the main specialty in which the consultant is contracted or recognised, it should be a valid code from the Main Specialty Codes table.

Treatment Function Code

This is the treatment function under which the patient is treated. It may be the same as the main specialty code or it may be a valid code from the Treatment Function Codes table.

First Attendance

An indicator for whether the patient was making a first or follow-up attendance and whether this was a face to face contact or telephone/telemedicine consultation.

Primary Procedure

The grouper application requires any interventions and procedures to be recorded as OPCS codes. Where no OPCS codes are recorded the grouper logic will assign an HRG based on the value in the first attendance.

Additional Procedures

Some of the grouper processing uses predefined combinations of codes to simplify the processing logic. These may be combinations of procedure codes referring to approaches used, or the site to which the procedure was applied. These should be recorded as OPCS codes. The grouper application will take all of the procedure codes into account when assigning HRGs.

3.6. File Preparation

The Reference Cost Grouper processes the data items shown above and if a meaningful HRG cannot be assigned a code of UZ01Z is returned signifying that the data is invalid for grouping.

In order to relate invalid records to source systems it may be valuable to add a unique key field to each record.

4. Emergency Medicine

4.1. Overview

Table 3 below shows the data items required by the Reference Cost Grouper.

Table 3

Grouper Name	Grouper Desc	Format	Validation	Desc / Derivation
INV_01	Accident And Emergency Investigation - First	an(2)	Null or valid code.	Valid national code
INV_02	Accident And Emergency Investigation - Second	an(2)	Null or valid code.	Valid national code
TREAT_01	Accident And Emergency Treatment - First	an(3)	Null or valid code.	Valid national code
TREAT_02	Accident And Emergency Treatment - Second	an(3)	Null or valid code.	Valid national code

Please note that validation permits records where all fields are null; such records should be avoided as default codes will be assumed by the grouper.

4.2. Record Preparation

The A&E Attendance CDS allows first and second investigations and first and second treatments to be recorded and so the application requires that all 4 fields are present and populated with either a valid national code or null.

Within the A&E Attendance CDS, the investigation fields can contain up to six alpha numeric characters. However, only the two leading characters are the national code, the last four positions being taken up by locally assigned values. These fields must be truncated to the two leftmost characters prior to submitting the data input file.

Within the A&E Attendance CDS, the treatment fields can contain up to six alpha numeric characters. Only the three leading characters are the national code, the last three positions being taken up by locally assigned values. These fields must be truncated to the three leftmost characters prior to submitting the data input file.

4.3. File Preparation

The Reference Cost Grouper processes the data items shown above and if a meaningful HRG cannot be assigned, a code of UZ01Z is returned signifying that the data is invalid for grouping.

In order to relate invalid records to source systems it may be valuable to add a unique key field to each record.

5. Adult Critical Care

5.1. Overview

Table 4 below shows the data items required by the Reference Cost Grouper.

Table 4

Field Name	Description	Format	Validation	Notes / Derivation
CCUF	Critical Care Unit Function Code	text	Not null	valid code for adult must be supplied
ARSD	Advanced Respiratory Support Days	number	0 - 99999	
BRSD	Basic Respiratory Support Days	number	0 - 99999	
ACSD	Advanced Cardiovascular Support Days	number	0 - 99999	
BCSD	Basic Cardiovascular Support Days	number	0 - 99999	
RSD	Renal Support Days	number	0 - 99999	
NSD	Neurological Support Days	number	0 - 99999	
DSD	Dermatological Support Days	number	0 - 99999	
CCL2D	Critical Care Level 2 Days	number	0 - 99999	
CCL3D	Critical Care Level 3 Days	number	0 - 99999	

N.B. Critical Care Start Date and Critical Care Discharge Date fields are no longer used in assigning activity to an HRG.

The inclusion of additional fields that help relate records to other systems may be valuable (e.g. to link the critical care record back to the appropriate admitted care episode). The following fields, where available, may be considered for inclusion:

Field Name	Description	Format	Validation	Notes / Derivation
PROCODET	Organisation Code	text	n/a	Local provider code possibly the NACS code (an(3))
PROVSPNO	Hospital Provider Spell Number	text	n/a	Local Spell Identifier
EPIORDER	Episode Number	number	n/a	

5.2. Record Preparation

A full description of how Adult Critical Care records should be prepared for Reference Costs submission is beyond the scope of this document. Further information is available in the 'Critical Care Minimum Dataset Training Pack' available from:

http://www.icservices.nhs.uk/casemix/pages/downloads/

5.3. File Preparation

The Reference Cost Grouper processes the data items shown above and if a meaningful HRG cannot be assigned a code of UZ01Z is returned signifying that the data is invalid for grouping.

In order to relate invalid records to source systems it may be valuable to add a unique key field to each record.